数理解析研究所講究録
842巻 1993年 76-100

Bayes risk and information inequalities

筑波大 数学 佐藤道一 (Michikazu Sato)
筑波大 数学 赤平昌文 (Masafumi Akahira)

§1 はじめに

Brown & Gajek (1990) (以下 [BG]) は、点推定で、適当な条件の下での Fisher 情報量を用いた Bayes 危険の下界を与えている。しかし、上記論文では事前密度が絶対連続であることを仮定しているので、例えば proper な一様分布の場合は適用できない。§2ではそのような場合について考察し、§§3~4では細かい注意を与え、§5では評価の比較をし、§§6~7ではミニマックス危険の下界を与える。

§2 Bayes 危険の下界について

確率変数 X はある σ 有限測度 ν に関する密度 p_θ を持つとする。θ は母数で、母数空間 Θ は $\mathbb{R}$ の区間とする。θ の点推定を考える。$a \in \mathbb{R}$ を推定値としたときの損失を

$$L(\theta, a) := (\theta - a)^2 \qquad (2.1)$$

で定める．推定量 $T=T(X)$ に対して

$$R(\theta, T) := E_\theta L(\theta, T)$$

とし、事前密度 g に対して

$$B(g, T) := \int R(\theta, T)\, g(\theta)\, d\theta$$

$$B(g) := \inf_T B(g, T)$$

とする．$B(g)$ を g の下での Bayes 危険 (Bayes risk) という．以上のことは §7 まで仮定する．但し、「標本の大きさが n」というときは X のかわりに $X_1, \cdots, X_n$ i.i.d. とする。なお (2.1) のかわりに

$$L(\theta, a) := m(\theta)(\theta - a)^2, \quad m(\theta) > 0$$

とした場合は、§4 までは $m(\theta)g(\theta)$ を $g(\theta)$ と考え直せば (2.1) の場合に帰着される。

次の (2a) ～ (2e) を仮定する．なお、「a.e. の θ」というときは Lebesgue 測度に関してである。また、「$[\theta_1, \theta_2]$ において C^n」というときは θ_1 では右微分、θ_2 では左微分を考えるので (2b) は矛盾した仮定ではないことに注意する。

(2a) $\theta_1, \theta_2 \in \overline{\Theta}$, $\theta_1 < \theta_2$ があって a.e. の $\theta \in (\theta_1, \theta_2)$ に対して Fisher 情報量

$$I(\theta) := E_\theta\left[\left\{\frac{\partial}{\partial\theta} \log p_\theta(X)\right\}^2\right]$$

が確定し、$V(\theta) := \dfrac{1}{I(\theta)}$ として $0 < V(\theta) \leqq \infty$

である。

(2b) 事前密度 g は $[\theta_1, \theta_2]$ において C^1 で $g(\theta)>0$, その他では $g(\theta)=0$ である.

(2c) 事前密度 g の下での Bayes 推定量

$$T_g(x) = \frac{\int_{\theta_1}^{\theta_2} \theta\, p_\theta(x)\, g(\theta)\, d\theta}{\int_{\theta_1}^{\theta_2} p_\theta(x)\, g(\theta)\, d\theta}$$

に対して、$\theta \mapsto E_\theta T_g \ (\theta \in (\theta_1, \theta_2))$ は $[\theta_1, \theta_2]$ における C^2 函数に拡張できる。

(2d) T_g に対して、a.e. の $\theta \in (\theta_1, \theta_2)$ で Cramér-Rao の不等式 [以下 CR と書く]

$$Var_\theta T_g \geqq V(\theta) \left\{\frac{d}{d\theta} E_\theta T_g\right\}^2$$

が成り立つ.

(2e) V_1 を $[\theta_1, \theta_2]$ において C^1 で

$$V_1(\theta) \leqq V(\theta) \qquad \text{a.e. } \theta \in (\theta_1, \theta_2)$$

$$0 < V_1(\theta) < \infty \qquad {}^\forall\theta \in [\theta_1, \theta_2]$$

となる函数とする。

<u>注意</u> $V=V_1$ として (2e) が満たされればそのときが最もよい評価となるが、下界を具体的に求めるのには一般には支障がある。

このとき、$b(\theta) := E_\theta T_g - \theta$ とすると、

$$R(\theta, T) \geqq V_1(\theta)\{1+b'(\theta)\}^2 + b^2(\theta)$$

$$B(g) = B(g, T_g) \geqq \int \{V_1(1+b')^2 + b^2\} g\, d\theta \tag{2.2}$$

となる．そこで、$[\theta_1, \theta_2]$におけるC^2函数$y = y(\theta)$に対して、

$$J(y) := \int \{V_1(1+b')^2 + b^2\} g\, d\theta$$

とおく．境界条件$y(\theta_1) = c_1$，$y(\theta_2) = c_2$を指定して$J(y)$を最小にしよう．ηを$[\theta_1, \theta_2]$におけるC^2函数で$\eta(\theta_1) = \eta(\theta_2) = 0$、$\eta \neq 0$とし、$\alpha$を定数として$J(y+\alpha\eta) - J(y)$を考えることにより、$y$で$J(y)$を最小にするための必要十分条件は、上の仮定を満たすすべてのηに対して

$$\int \{V_1 \eta'(1+y') + \eta y\} g\, d\theta = 0 \tag{2.3}$$

となることである。部分積分により、(2.3)は

$$\int \eta \{y''(V_1 g) + (1+y')(V_1 g)' - g y\} d\theta = 0 \tag{2.4}$$

と書きかえられる。これが仮定を満たすすべてのηに対して成り立つための必要十分条件は、yが微分方程式

$$y''(V_1 g) + (1+y')(V_1 g)' - g y = 0$$

の解となることである。この微分方程式は

$$y'' + \tilde{g} y - \frac{1}{V_1} y + \tilde{g} = 0 \tag{2.5}$$

$$但し \quad \tilde{g} := (\log V_1 g)'$$

と書きかえられる。

<u>補題2.1</u>　微分方程式（2.5）は任意の境界条件に対してただ1つの解を持ち、解はC^2である。

<u>証明</u>　ただ1つ持つことを示すには、

$$y'' + \tilde{g} y' - \frac{1}{V_1} y = 0, \quad y(\theta_1) = y(\theta_2) = 0 \tag{2.6}$$

の解が$y \equiv 0$に限ることを示せばよい［例えば、草野（1971）p. 49］。それには、否として最大（小）値での符号を考えて矛盾を導けばよい。C^2は（2.5）のy''以外の項を移項すればよい。 □

<u>定理2.1</u>　上記の仮定の下で、微分方程式（2.5）の一般解を$y = y_0 + Ay_1 + By_2$（y_1, y_2は一次独立）の形に表すと、$J(y)$は、yを$[\theta_1, \theta_2]$におけるC^2函数全体を動かしたときに最小値J_0を持ち、

$$B(g) \geqq J_0 = \frac{a_0 b_0 c_0 + 2 f_0 g_0 h_0 - a_0 f_0^2 - b_0 g_0^2 - c_0 h_0^2}{a_0 b_0 - h_0^2}$$

となる。但し、

$$a_0 := \int (V_1 y_1'^2 + y_1^2) g \, d\theta$$
$$b_0 := \int (V_1 y_2'^2 + y_2^2) g \, d\theta$$
$$c_0 := \int \{V_1 (1 + y_0')^2 + y_0^2\} g \, d\theta$$

$$f_0 := \int \{V_1(1+y_0')y_2' + y_0 y_2\} g\, d\theta$$
$$g_0 := \int \{V_1(1+y_0')y_1' + y_0 y_1\} g\, d\theta$$
$$h_0 := \int (V_1 y_1' y_2' + y_1 y_2)\, g\, d\theta$$

とする。

<u>証明</u>　$J(y_0 + A y_1 + B y_2)$ は A, B の２次式であり、

$$\langle y_1, y_2 \rangle := \int (V_1 y_1' y_2' + y_1 y_2)\, g\, d\theta$$

にCauchy-Schwarzの不等式を適用して $h_0^2 < a_0 b_0$ を得る。そこで $\begin{pmatrix} a_0 & h_0 \\ h_0 & b_0 \end{pmatrix}$ の固有値を λ_1, λ_2 とすると、$\lambda_1, \lambda_2 > 0$ であり、座標変換を考え［例えば、石谷(1981)］、最小値は

$$J_0 = \frac{\begin{vmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & J_0 \end{vmatrix}}{\begin{vmatrix} \lambda_1 & \\ & \lambda_2 \end{vmatrix}} = \frac{\begin{vmatrix} a_0 & h_0 & g_0 \\ h_0 & b_0 & f_0 \\ g_0 & f_0 & c_0 \end{vmatrix}}{\begin{vmatrix} a_0 & h_0 \\ h_0 & b_0 \end{vmatrix}}$$

となる。 □

<u>例2.1</u>　$\theta_1 = \theta_0 - \delta$, $\theta_2 = \theta_0 + \delta$ $(\delta > 0)$ とする。$\lambda \in \mathbb{R}$ として、事前密度 g を

$$g(\theta) := \frac{e^{\lambda(\theta - \theta_0)}}{\varphi(\lambda)} \qquad (|\theta - \theta_0| < \delta)$$

で定める。但し、

$$\varphi(t) := \int_{-\delta}^{\delta} e^{t\theta} d\theta = \begin{cases} \dfrac{e^{\delta t} - e^{-\delta t}}{t} & (t \neq 0) \\ 2\delta & (t = 0) \end{cases}$$

とする。(2a)(2c)(2d)と、次の(2f)

(2f)　$v_* := \inf_{\theta_1<\theta<\theta_2} V(\theta)$, 　$0 < v_* < \infty$, 　$V_1 \equiv v_*$

を仮定する。このとき、

$$B(g) \geqq J_0 = v_*(v_*\lambda^2+1) - \frac{v_*\sum_{j=1}^{2}\nu_j^2(v_*\nu_j^2+1)\varphi^2(\nu_j)}{\rho^2\varphi(\rho)\varphi(\lambda)}$$

但し　$\rho := \sqrt{\lambda^2+\frac{4}{v_*}}$, 　$\nu_1 := \frac{-\lambda+\rho}{2}$　$\nu_2 := \frac{-\lambda-\rho}{2}$

となる。特に $\lambda=0$ のときは事前分布は一様分布 $U(\theta_1, \theta_2)$ となり、

$$B(g) \geqq J_0 = v_*\left(1-\frac{\sqrt{v_*}\,\kappa}{\delta}\right) \qquad (2.7)$$

但し　$\gamma = e^{\frac{\delta}{\sqrt{v_*}}}$, 　$\kappa = \frac{\gamma-\gamma^{-1}}{\gamma+\gamma^{-1}}$ 　$(0<\kappa<1)$

となる。なお、(2f) のかわりに

(2g)　$0 < v_* \leqq \inf_{\theta_1<\theta<\theta_2} V(\theta)$, 　$v_* < \infty$, 　$V_1 \equiv v_*$

としても成り立つ。

証明　　まず, $\theta_0 = 0$ としてよい。

$y_0 \equiv v_*\lambda$, 　$y_1 = e^{\nu_1\theta}$, 　$y_2 = e^{\nu_2\theta}$ であり、定理2.1の式に代入すればよい。 $\square$

なお、g を [BG] の仮定を満たすもので近似して極限をとるという方法では、通常は(常に?) [BG] の記号で $D\to\infty$ となってしまい不都合が生ずる。

§3　下界達成について

§2で、$B(g)=J_0$ とはならないのがむしろ普通である。まず正則条件を述べる。T_g は (2c) の記号である。

(3a)　$p_\theta(x)>0 \quad {}^\forall x \quad {}^\forall\theta\in(\theta_1,\theta_2)$

(3b)　任意の x を固定したとき、$\theta\mapsto p_\theta(x)$ は (θ_1,θ_2) で C^1 である。

(3c)　$\int p_\theta(x)\nu(dx)$ は $\theta\in(\theta_1,\theta_2)$ において θ について積分記号内で微分可能である。

(3d)　$\int T_g(x)p_\theta(x)\nu(dx)$ は $\theta\in(\theta_1,\theta_2)$ において θ について積分記号内で微分可能である。

(3e)　V_1 は $[\theta_1,\theta_2]$ において C^1 で、(2a) の記号で

$$0<V_1(\theta)\leqq V(\theta)<\infty \qquad {}^\forall\theta\in(\theta_1,\theta_2)$$

である。

(3f)　$\theta\mapsto Var_\theta T_g$ は (θ_1,θ_2) で連続である。

<u>定理3.1</u>　(2b), (2c), (3a)〜(3f) [従って、(2a)〜(2e)] を仮定する。このとき、$B(g)=J_0$ なら、$\nu(K)=0$ となる K があって、

$$p_\theta(x)=c(\theta)h(x)e^{q(\theta)T_g(x)}$$
$$(x\notin K,\ \theta\in(\theta_1,\theta_2))$$

ここで、$c,h>0$, q は狭義単調, c,q は C^1

と表されなければならない.

注意　T_g は本質的に有界である.

証明　$B(g)=J_0$ ならbは (2.5) の解でなければならず、$E_\theta T_g$ が定数なら矛盾が生ずる. あとは Wijsman (1973)（以下 [W]）による. □

なお、[W] の正則条件をゆるめた場合については Joshi (1976) が述べている.

定理3.2　ν−a.e. で定数でない推定量 T_0 があって、(2c), (3a)〜(3e)（但し、(2c)(3d) は T_g のかわりに T_0 とする）が成り立ち、T_0 に対して任意の $\theta\in(\theta_1,\theta_2)$ で CR の等号が成立すると仮定する. このとき、事前密度 g に対して (2b)(3d) が成り立ち、$B(g)=J_0$ となるならば、T_0 は本質的に有界で、V_1, g, T_g は次の形に表されなければならない.

$$V_1(\theta)=V(\theta)\qquad(\theta_1<\theta<\theta_2)\tag{3.1}$$

$$g(\theta)=cI\exp\int^{\theta}\frac{(\alpha m-\theta+\beta)I-\alpha m''}{\alpha m'}d\theta\quad(\theta_1<\theta<\theta_2)\tag{3.2}$$

$$T_g=\alpha T_0+\beta\qquad \nu-\text{a.e.}\tag{3.3}$$

但し、$m(\theta):=E_\theta T_0$ とし、積分は「原始函数をしっとって固定する」という意味であり、$\alpha(\neq0)$, β, c は定数である.

更に、区間 $K\subset\overline{\Theta}$ があって、任意の $\theta_1,\theta_2\in K\ (\theta_1<\theta_2)$

に対して仮定が成り立つとき、α、$\beta\in\mathbb{R}$ $(\alpha\neq 0)$, $\theta_1\in K$を固定して$\theta_2\in K\ (\theta_2>\theta_1)$を動かして$g$を(3.2)で定めたとき、(3d)を満たし$B(g)=J_0$となる$\theta_2$は高々1つしかない。$\alpha$, β, θ_2を固定してθ_1を動かしても同様である。

証明　gが仮定を満たすと、[W]より(3.3)と表され、T_0は本質的に有界で、$m'(\theta)\neq 0$はCR達成と(3a)とT_0がν-a.e.で定数ではないことから得られ、(3.1)はこのこととJ_0達成、(2b)(3b)(3e)とFatouの補題から得られ、$\alpha\neq 0$は定理3.1の証明の過程から得られる。(3.3)からbを求めて(2.5)に代入してgについて解けば(3.2)が得られる。後半のθ_2の一意性を示そう。2つあったとし、θ_2^*, $\theta_2^{**}\in K\ (\theta_1<\theta_2^*<\theta_2^{**})$とすると、

$$g_1(\theta):=I\exp\int^{\theta}\frac{(\alpha m-\theta+\beta)I-\alpha m''}{\alpha m'}d\theta$$

として、

$$\alpha T_0(x)+\beta=\frac{\int_{\theta_1}^{\theta_2^*}\theta p_\theta(x)g_1(\theta)d\theta}{\int_{\theta_1}^{\theta_2^*}p_\theta(x)g_1(\theta)d\theta}=\frac{\int_{\theta_1}^{\theta_2^{**}}\theta p_\theta(x)g_1(\theta)d\theta}{\int_{\theta_1}^{\theta_2^{**}}p_\theta(x)g_1(\theta)d\theta}$$

がν-a.e.のx、従って少なくとも1つのx（それを固定する）で成り立つので、上式の値をaとして、

$$F(\theta_2):=\int_{\theta_1}^{\theta_2}(\theta-a)p_\theta(x)g_1(\theta)d\theta$$

の増減を考えれば矛盾が得られる。　□

例3.1　$X\sim N(\theta,1)$のとき、(2b)を満たす任意のg

に対して、$B(g) > J_0$ である。 □

例3.2 $X \sim B(n, \theta)$ のとき [(2a) より $0 < \theta_1 < \theta_2 < 1$] 定理3.2で $T_0 = X$ とし、$B(g) = J_0$ ならば

$$g(\theta) = c\theta^{\frac{\beta}{\alpha}-1}(1-\theta)^{\frac{-\beta+1}{\alpha}-n-1} \quad (\theta_1 < \theta < \theta_2)$$

でなければならない。 □

なお、実際に $B(g) = J_0$ となる例があるかどうかは不明である。

§4 仮定が満たされない場合について

仮定(2b)より、g は θ_1, θ_2 で不連続である。これが満たされない場合は微分方程式(2.5)が任意の境界値に対して解を持つとは限らない。例として、$\theta_1 = 0$, $\theta_2 = \delta > 0$ とし、

$$g(\theta) = c\theta^{\lambda} \quad (0 \leqq \theta \leqq \delta)$$

の場合を考える。(2a)(2c)(2d)(2f) を仮定する。$\lambda < 0$ なら $\theta \downarrow 0$ のとき $g(\theta) \to \infty$ となるので(2.5)に持ち込む以前の部分積分で問題があるので除外する。$\lambda = 0$ なら例2.1である。$\lambda > 0$ の場合を考える。(2.5)は $0 < \theta \leqq \delta$ で考え、$y(0) = \lim_{\theta \downarrow 0} y(\theta)$ とする。解の1つは

$$y_0 = -\sum_{j=0}^{\infty} \frac{\theta^{2j+1}}{\left\langle {2j+\lambda \atop \lambda} \right\rangle \left\langle {2j+1 \atop \lambda} \right\rangle}$$

但し、$\left\langle {n \atop 0} \right\rangle = 1$, $\left\langle {n \atop j} \right\rangle = \underbrace{n(n-2)(n-4)\cdots(n-2j+2)}_{j \text{個}}$

となる。この収束半径は∞である。あとは斉次方程式を吟味すればよい。それにはFrobeniusの理論[例えば、笠原(1982), 草野(1971)]を用いる。$\lambda \geqq 1$ なる任意の c_1 に対し、ただ1つの c_2 に対してのみ解を持つ。$0<\lambda<1$ なる任意の境界値に対して解を持つ。この解に対し、(2.5)に持ち込むまでの計算が正しいことを別途確かめる必要がある。但し $J(y)$ の y の動かす範囲は $(0, \delta]$ で C^2 で 0 で右連続と変える、η も同様に変える。吟味しなければならないのは、y が解のとき $J(y)<\infty$ であることと、(2.3)から部分積分によって(2.4)に持ち込む際

$$\lim_{\theta \downarrow 0} \eta (1+y')(V_1 g) = 0$$

が成り立つことである。それは今の場合は成り立つ。

§5 評価の比較

$\theta_1 = \theta_0 - \delta$, $\theta_2 = \theta_0 + \delta$ $(\delta > 0)$ とし、事前密度を

$$g(\theta) := \frac{1}{\delta} \cos^2 \frac{\pi}{2\delta}(\theta - \theta_0) \qquad (|\theta - \theta_0| < \delta)$$

で定める。この事前分布を $Cos^2(\theta_0, \delta)$ で表すことにする。

(2a)(2d)(2f)[もしくは(2g)]と次の(5a)

(5a) (2c)の記号で $\theta \mapsto E_\theta T_g$ は $[\theta_1, \theta_2]$ における絶対連続函数に拡張できる。

を仮定する。このとき、Borovkov & Sakanienko (1980)(以下[B,S])、[BG]より、

$$B(g) > v_*\left(1+\frac{\pi^2 v_*}{\delta^2}\right)^{-1} \qquad (5.1)$$

となる。なお[BG]では(5a)のかわりに微分可能としてあるが、これには疑問の余地がある。一般に f が $[x_1, x_2]$ で C^1 なら絶対連続だが、微分可能でも絶対連続とは限らない[吉田(1965) pp. 232−233]。なお今の例で、V は絶対連続とは限らないが、V_1 が絶対連続なので[BG]の系2.3と注意2.5を用いるのに支障はない。注意2.8も同様である。

一般に、

$$R^*(T) := \sup_\theta R(\theta, T), \quad r^* := \inf_T R^*(T)$$

とする。r^* をミニマックス危険(minimax risk)といい、$R^*(T_0) = r^* < \infty$ のとき T_0 はミニマックスであるという。一般に、$r^* \geqq B(g)$ が成り立つので、§2及び[BS]、[BG]の $B(g)$ の評価は r^* の評価と考えることができる。そこで例2.1の(2.7)と今の(5.1)の右辺を r^* の評価と

考え、この良し悪しを考えると、v_* が十分小さいときは (5.1) の方が良い評価であり、v_* が十分大きいときは (2.7) の方が良い評価である $\left(x = \frac{\delta}{\sqrt{v_*}}\right.$ とし、

$$x^3\left\{\frac{(2.7)\text{ の右辺}}{(5.1)\text{ の右辺}} - 1\right\} \to -\pi^3 < 0 \quad (x \uparrow \infty)$$

$$(e^{2x}+1)\left\{\frac{(2.7)\text{ の右辺}}{(5.1)\text{ の右辺}} - 1\right\} \to \frac{2\pi^2}{3} - 2 > 0 \quad (x \downarrow 0)$$

による)。なお、標本の大きさが n のときを考えると (r^* を r_n^* と書く)、v_* は $\frac{v_*}{n}$ で置き換えられるので、今のことより漸近的には (5.1) の方が良い評価で、

$$n r_n^* > v_*\left(1 + \frac{\pi^2 v_*}{\delta^2 n}\right)^{-1} \qquad (5.2)$$

となるが、固定した n に対しては (2.7) の方が良い評価となることもある。なお、後述のように、(5.2) も一般には漸近的によい評価とはいえない。

§6　ミニマックス危険の下界について (固定標本)

r^* の良い評価を得るには、一般には事前分布の族を考えて極限操作をする必要がある。また §5 のように v_* を用いるのではなく、

$$v^* := \sup V(\theta)$$

但し sup は $V(\theta)$ が確定する範囲の θ を動かす．

を用いるべきである。まず正則条件を述べる。

(6a)　θ_0 があって $(\theta_0, \infty) \subset \Theta$ であり，a.e. の $\theta > \theta_0$ に対して (2a) の $V(\theta)$ が確定する。

(6b)　任意の有界推定量 T に対して，$\theta \mapsto E_\theta T$ は (θ_0, ∞) に含まれる任意の有界閉区間上で絶対連続で，a.e. の $\theta > \theta_0$ で CR が成り立つ。

定理 6.1　(6a)(6b) を仮定する。このとき，

$$r^* \geqq \liminf_{\theta \to \infty} V(\theta) \quad [=: v_\infty]$$

である。

証明　$v_\infty > 0$ としてよい．$0 < M < \infty$ を固定し，十分大きい $k \in N$ に対して，

$$v_*(k) := \min\{\inf_{\theta > k} V(\theta),\ M\}$$

とすると，v_* は単調増加で，

$$\lim_{k \to \infty} v_*(k) = \min\{v_\infty,\ M\}$$

であり，事前分布を $Cos^2(2k, k)$ として，(5.1) より

$$r^* > v_*(k)\left(1 + \frac{\pi^2 v_*(k)}{k^2}\right)^{-1}$$

だから，$k \to \infty$ として，

$$r^* \geqq \min\{v_\infty,\ M\}$$

となり，これが任意の $0 < M < \infty$ で成り立つので $r^* \geqq v_\infty$

である。 □

<u>注意</u> $v_\infty = v^*$ でないときはあまり良い評価ではない。なお、$\theta \to \infty$ のかわりに $\theta \to -\infty$ を考えても同様である。

許容性については、次のことが成り立つ。まず正則条件を述べる。

(6c) $\Theta = \mathbb{R}$ で、任意の θ に対して (2a) の $V(\theta)$ が確定し、$0 < V(\theta) < \infty$ である。

(6d) $\lim_{\theta \to \infty} V(\theta)$, $\lim_{\theta \to -\infty} V(\theta)$ が有限確定である。

(6e) $E_\theta T^2 < \infty \ \forall\theta$ を満たす任意の推定量 T に対して $E_\theta T$ は θ に関して微分可能で CR が成り立つ。

<u>定理 6.2</u> (6c)〜(6e) を仮定する。このとき、θ の不偏推定量 T_0 が CR で任意の θ で等号が成り立てば、T_0 は許容的である。

<u>証明</u> 鍋谷 (1978) 定理 6.2.1 と同様である。なお、数列 $\{\theta_i\}$ は b に対して $[i, i+1]$ で平均値の定理を用いれば得られる。 □

<u>例 6.1</u> 定理 6.1 で特に $v_\infty = v^* < \infty$ で、θ の不偏推定量 T_0 が CR で任意の θ で等号が成り立てば T_0 はミニマックスである。例えば $X \sim N(\theta, 1)$ $(\Theta = \mathbb{R})$ のとき X はミニマックスであり、定理 6.2 より許容的である。なお、$\Theta = (0, \infty)$ としてもミニマックスであるが、この場合は

$\max\{X, 0\}$ で改良されるので非許容的である。 $\square$

§7　ミニマックス危険の下界について（漸近論）

§6では固定標本で考えたが、Ⓗが有界のときは、後述の例7.1のように固定標本ではうまくいかない。この場合、正則条件の下で漸近的に良い評価を与えることができる。

$T_n = T_n(X_1, \cdots, X_n)$ を標本の大きさが n のときの推定量とする。

$$\lim_{n\to\infty} \frac{r_n^*}{R^*(T_n)} = 1 \qquad \left(\text{但し} \frac{0}{0} = 1 \text{とする。}\right)$$

が成り立つとき、T_n は（厳密には $\{T_n\}_{n=1}^{\infty}$ は）漸近的にミニマックスであるという。

正則条件を述べる。

(7a)　$-\infty \leqq \theta_0 < \theta_1 < \infty$ があって $(\theta_0, \theta_1) \subset$ Ⓗ で、a.e. の $\theta \in (\theta_0, \theta_1)$ で $V(\theta)$ が確定して $v^* = \lim_{\theta \downarrow \theta_0} V(\theta)$ である。

(7b)　任意の有界推定量 T_n に対して、$\theta \mapsto E_\theta T_n$ は (θ_0, θ_1) に含まれる任意の有界閉区間上で絶対連続で、a.e. の $\theta \in (\theta_0, \theta_1)$ で CR が成り立つ。

<u>定理7.1</u>　(7a), (7b) を仮定する。このとき、

$$\liminf_{n\to\infty} n r_n^* \geqq v^*$$

である。

<u>証明</u> $\theta_0 = -\infty$ のときは定理6.1（と注意）より

$n r_n^* \geq v^*$ だから当然成り立つ。否のときを示せばよい。

$\theta_0 = 0$、$v^* > 0$ としてよい。$0 < M < \infty$ を固定して、

$$V_+(\delta) := \min\left\{\inf_{0<\theta<\delta} V(\theta),\ M\right\}$$

とすると、(7a) より

$$V_+(\delta) \to \min\{v^*, M\} \qquad (\delta \downarrow 0)$$

だから、十分小さい $\delta > 0$ に対して、事前分布を

$Co_2^2(2\delta, \delta)$ として、(5.2) より

$$n r_n^* > V_+(3\delta)\left(1 + \frac{\pi^2 V_+(3\delta)}{\delta^2 n}\right)^{-1}$$

だから、辺々の $\liminf_{n\to\infty}$ をとり、次いで $\delta \downarrow 0$ とし、$M \uparrow \infty$

とすればよい。 □

<u>注意</u> (7a) で $\theta \downarrow \theta_0$ のかわりに $\theta \uparrow \theta_0$ を考えても

同様である。また、今の定理から (5.2) は $v_* < v^*$ のと

きは漸近的に良くない評価であることがわかる。なお、今の

定理は n を固定して意味がない。十分大きい n を固定して意

味を持ち、しかも漸近的に良い評価は次の定理で得られる。

<u>定理7.2</u> (I) 次のことを仮定する。

(7c) $\quad 0 < v^* < \infty$

(7d) $\quad \theta_0 \in \Theta^\circ$ で、δ_0, d, $k > 0$ があって、a.e.の

$\theta \in (-\delta_0, \delta_0)$ に対して

$$V(\theta_0+\theta) \geqq v^* - k|\theta|^d$$

が成り立つ.

(7e) 任意の有界推定量T_nに対して、$\theta \mapsto E_\theta T_n$は $(\theta_0-\delta_0, \theta_0+\delta_0)$ に含まれる任意の閉区間上で絶対連続で、a.e. の $\theta \in (\theta_0-\delta_0, \theta_0+\delta_0)$ で CR が成り立つ.

このとき、

$$C := (2\pi^2 d^{-1})^{\frac{d}{d+2}} k^{\frac{2}{d+2}} (d+1)(v^*)^{\frac{d-2}{d+2}}$$

(nによらない) として、

$$\frac{n r_n^*}{v^*} > 1 - C n^{-\frac{d}{d+2}} \qquad (7.1)$$

但し、$n > 2(\pi v^*)^2 \delta_0^{-d-2} (kd)^{-1}$

が成り立つ. 特に$d=2$のとき、

$$\frac{n r_n^*}{v^*} > 1 - 3\pi\sqrt{\frac{k}{n}} \qquad (7.2)$$

但し、$n > (\pi v^*)^2 \delta_0^{-4} k^{-1}$

である.

(II) (7c) と、次のことを仮定する.

(7f) $\theta_0 \in \overline{\Theta}$ で、$\delta_0, d, k > 0$ があって、a.e. の $\theta \in (0, \delta_0)$ に対して

$$V(\theta_0+\theta) \geqq v^* - k\theta^d$$

が成り立つ.

(7g) 任意の有界推定量T_nに対して、$\theta \mapsto E_\theta T_n$は$(\theta_0, \theta_0+\delta_0)$に含まれる任意の閉区間上で絶対連続で、a.e.の$\theta \in (\theta_0, \theta_0+\delta_0)$でCRが成り立つ。

このとき、

$$D := 2^{\frac{3d}{d+2}} (\pi^2 d^{-1})^{\frac{d}{d+2}} k^{\frac{2}{d+2}} (d+1)(v^*)^{\frac{d-2}{d+2}}$$

(nによらない)として、

$$\frac{n r_n^*}{v^*} > 1 - D n^{-\frac{d}{d+2}} \qquad (7.3)$$

但し、$n > 2^{d+3}(\pi v^*)^2 \delta_0^{-d-2}(kd)^{-1}$

が成り立つ。

証明 $\theta_0 = 0$としてよい.

(I) $0 < \delta < \delta_0$に対して、事前分布を$\mathrm{Cos}^2(0, \delta)$とすると、標本の大きさが1のとき、(5.1)より、$v^* > k\delta^d$のとき、

$$r^* > (v^* - k\delta^d)\left\{1 + \frac{\pi^2}{\delta^2}(v^* - k\delta^d)\right\}^{-1}$$

だから、$\frac{1}{1+x} > 1-x \ (x>0)$を用いて、

$$\frac{r^*}{v^*} > 1 - \frac{k\delta^d}{v^*} - \frac{\pi^2 v^*}{\delta^2}$$

となる．これは $v^* \leqq k\delta^d$ のときを含めて成り立つ．標本の大きさが n のときは，r^* を r_n^*，v^* を $\frac{v^*}{n}$，k を $\frac{k}{n}$ として，

$$\frac{n r_n^*}{v^*} > 1 - \frac{k\delta^d}{v^*} - \frac{\pi^2 v^*}{\delta^2 n}$$

となる．$\delta = cn^{-\lambda}$ $(c, \lambda > 0)$ とすると，十分大きい n に対して，$0 < \delta < \delta_0$ で，特に $\lambda = \frac{1}{d+2}$，次いで $c = \{2(\pi v^*)^2 (kd)^{-1}\}^{\frac{1}{d+2}}$ として（これは漸近的に評価を良くするためである．），(7.1) 式を得る．n の範囲は $\delta < \delta_0$ を n について解けば得られる．

(II) 事前分布を $\mathrm{Cos}^2\left(\frac{1+\varepsilon}{2}\delta, \frac{1-\varepsilon}{2}\delta\right)$ $(0 < \delta < \delta_0,\ \varepsilon > 0$ は十分小$)$ として，

$$r^* > (v^* - k\delta^d)\left\{1 + \frac{\pi^2}{\left(\frac{1-\varepsilon}{2}\delta\right)^2}(v^* - k\delta^d)\right\}^{-1}$$

だから $\varepsilon \downarrow 0$ とし，(I) と同様の計算で，

$$\frac{r^*}{v^*} > 1 - \frac{2^d k}{v^*}\left(\frac{\delta}{2}\right)^d - \frac{\pi^2 v^*}{\left(\frac{\delta}{2}\right)^2}$$

となるから，(I) で δ を $\frac{\delta}{2}$，k を $2^d k$ として (7.3) 式を得る．n の範囲は $\frac{\delta}{2} = cn^{-\lambda}$ として $\delta < \delta_0$ であればよいから，δ_0 を $\frac{\delta_0}{2}$ として得られる． □

この定理で，d は仮定を満たす範囲で大きくとった方が漸

近的に評価は良くなり、dを固定するとkは仮定を満たす範囲で小さくとった方が漸近的に評価は良くなる。δ_0 は (7.1)〜(7.3) には現れずnの範囲に現れるだけであり、d, kを固定すればδ_0 は仮定を満たす範囲で大きくとった方がnの範囲を広くできるが、一般には小さくとった方がkをとり直して漸近的に評価を良くできる。なお、Vの最大値を与えるθの集合が区間を含むときは、その区間を $(\theta_0-\delta_0, \theta_0+\delta_0)$ として、(5.2) より

$$\frac{n r_n^*}{v^*} > \left(1+\frac{\pi^2 v^*}{\delta_0^2 n}\right)^{-1} > 1-\frac{\pi^2 v^*}{\delta_0^2 n}$$

とした方が漸近的に良い評価である。

なお、Vがθ_0 で最大で、$[\theta_0-\delta_0, \theta_0+\delta_0]$ でC^d (dは偶数) で、$V'(\theta_0)=V''(\theta_0)=\cdots=V^{(d-1)}(\theta_0)=0$, $V^{(d)}(\theta_0)\neq 0$ のときは、$|\theta|<\delta_0$ に対して $0<\eta<1$ があって

$$V(\theta_0+\theta)=v^*+V^{(d)}(\eta\theta)\theta^d \geqq v^*-k|\theta|^d$$

$$\text{但し } k:=-\inf_{|\theta|<\delta_0} V^{(d)}(\theta_0+\theta)>0$$

となる。$\delta_0\downarrow 0$ とすると $k\to -V^{(d)}(\theta_0)$ となるので

$$\liminf_{n\to\infty} n^{\frac{d}{d+2}}\left(1-\frac{n r_n^*}{v^*}\right)$$

$$\leqq (2\pi^2 d^{-1})^{\frac{d}{d+2}} |V^{(d)}(\theta_0)|^{\frac{d}{d+2}} (d+1)(v^*)^{\frac{d-2}{d+2}}$$

特に$d=2$のときは

$$\limsup_{n\to\infty}\sqrt{n}\left(1-\frac{n r_n^*}{v^*}\right)\leqq 3\pi\sqrt{|V''(\theta_0)|}$$

となるが、これはnを固定して意味のない式である。

<u>例7.1</u> *Bernoulli* 試行、すなわち

$X_1, X_2, \cdots \sim B(1, \theta)$ i.i.d. $0<\theta<1$ の場合

$$V(\theta)=\theta(1-\theta), \quad V\left(\frac{1}{2}+\theta\right)=\frac{1}{4}-\theta^2$$

だから、$v^*=\frac{1}{4}$, $\theta_0=\delta_0=\frac{1}{2}$, $d=2$, $k=1$として

定理7.2 (I)を用いると、

$$\frac{n r_n^*}{v^*}>1-\frac{3\pi}{\sqrt{n}} \qquad (n\geqq 10)$$

となる. もっとも自明でない評価 (すなわち、右辺>0) と

なるためには$n\geqq 89$である. θの不偏推定量$U_n=\frac{1}{n}\sum_{j=1}^{n}X_j$

はCRで等号が成立するので、定理7.2 (もしくは7.1) より漸近的にミニマックスである. しかし、U_nはミニマックスではない. 正確には、

$$T_n=\frac{\sum_{j=1}^{n}X_j+\frac{\sqrt{n}}{2}}{n+\sqrt{n}}$$

がミニマックスで

$$r_n^*=R^*(T_n)=R(\theta, T_n)=\frac{1}{4(\sqrt{n}+1)^2}$$

となり [例えば、鍋谷 (1978) p. 155]

$$\frac{n r_n^*}{v^*} = 1 - \frac{2\sqrt{n}+1}{n+2\sqrt{n}+1} = 1 - a_n \quad 但し\ a_n \sim \frac{2}{\sqrt{n}}$$

と表される．従って、定理7.2（I）で$d=2$として仮定が満たされる場合、

$$\frac{n r_n^*}{v^*} \geqq 1 - C n^{-\rho} \qquad (nは十分大)$$

の形の評価を考えると、この定理の$\rho = \frac{1}{2}$を改良することはできず、改良の余地があるとすればCである。 □

参考文献

Borovkov, A. A. & Sakhanienko, A. U. (1980) On estimate of the expected quadratic risk, Probab. Math. Statist. 1, 185–195. (In Russian.)

Brown, L. D. & Gajek, L. (1990) Information inequalities for the Bayes risk. Ann. Statist. 18. 1578–1594.

石谷茂 (1981) 新数学対話4 矢線ベクトル．現代数学社．

Joshi, V. M. (1976) On the attainment of the Cramér-Rao lower bound, Ann. Statist. 5. 998–1002.

笠原晧司 (1982) 微分方程式の基礎．朝倉書店．

草野尚 (1971) 境界値問題入門．朝倉書店．

鍋谷清治 (1978) 数理統計学. 共立出版.

吉田洋一 (1965) ルベグ積分入門. 培風館.

Wijsman, R. A. (1973) On the attainment of the Cramér-Rao lower bound, Ann. Statist. 1. 538-542.